

Lung Cancer Survival and Mortality Analyses of Guam Cancer Registry Data, 2000-2009  
Final Report

by  
Michael B. Ehlert, Ph.D.  
University of Guam

Author Note

Please address correspondence to the author at (email: [research@ehlertweb.net](mailto:research@ehlertweb.net) or [mbehlert@ugam.uog.edu](mailto:mbehlert@ugam.uog.edu)) or (phone: 671-688-5590 or 671-735-2870).

The project was funded by Guam Department of Public Health and Social Services through the Guam Comprehensive Cancer Control Coalition Data and Research Action Team (DRAT), R. Bordallo, Team Leader.

Special acknowledgement to Jasmin Benavente, UOG Psychology Undergraduate, for assistance with data analyses; Renata Bordallo, Guam Cancer Registry and DRAT Team Leader; and to those DRAT Team members who provided comments and suggestions.

## Lung Cancer Survival and Mortality Analyses of Guam Data, 2000-2009

The Guam Cancer Facts and Figures, 2003-2007 (2009) reported 1,580 new cancer cases (*incidence*) and 720 deaths (*mortality*) due to cancer. Incidence and mortality data are fundamental to understanding cancer and its impact. Another key metric, however, is cancer survival. Cancer survival analysis is a means to understand the effect of cancer by assessing how long individuals live after a cancer diagnosis. To advance our understanding of the burden of cancer on Guam, the Guam Department of Public Health and Social Services (DPHSS) solicited a cancer survival analysis through the Guam Comprehensive Cancer Control Coalition's Data and Research Action Team (DRAT).

### Cancer Survival Concepts

Survival analyses require a starting event, the date of a first cancer diagnosis in this study, and a terminating event, a cancer patient's vital status on a specific date, 31 December 2009 in this study. The *survival rate* provides a numerical summary of targeted cases on a specified date, typically 5 years. The *survival curve* plots the survival rate for particular durations since diagnosis (see Compton et al., 2012).

Researchers have developed various methods to conduct survival analyses, each with specific strengths and weaknesses (for reviews see Parkin & Hakulinen, 1991, and Compton et al., 2012). The *direct method* calculates the proportion of cases alive after a specified period of time. But the direct method excludes from analysis some data that may be useful (i.e., cases who survived part of the specified interval). Two more approaches include the *actuarial* method (also called *life-table*) and the *Kaplan-Meier* method. The actuarial method distinguishes between confirmed deaths during a specified period and cases that were last seen alive but may have deceased since last follow up (see Parkin & Hakulinen, 1991, Table 2 for computation example). The Kaplan-Meier method calculates survival rate at whatever interval the date of death is recorded (usually month but it could be even to the day) and assumes that *withdrawals* (lost cases that cannot be located) survive to the end of the interval. Kaplan-Meier provides the most accurate estimate of a survival curve if all necessary data are available (Parkin & Hakulinen, 1991; see Table 5 for a computation example).

The three methods described provide different estimates of the *observed* survival rate. Adjustments to observed survival are possible and allow better comparison across populations. The usual adjustments include age and cause of death. Populations have different mean ages making it likely that cancer incidences differ depending on the relative proportion of younger and older persons existing in the population(s) of interest (called *age-adjusted*). For example, a population with more children should report less cancer than one with a higher proportion of older citizens. A second important factor considers the specific cause of death. Adjustment for the actual cause of death focuses the analysis on deaths due to cancer and excludes cancer patients who die from other causes (called *relative survival*; Cho, Howlader, Mariotto, & Cronin, 2011). Neither of these adjustments is incorporated in this study and, therefore, comparison to other populations must be done cautiously.

A *mortality* analysis is related to a survival analysis but focuses exclusively on those cases whose vital status is 'deceased' due to cancer. This type of analysis investigates how

long those patients whose deaths are attributable to cancer survive after the diagnosis, instead of including *all* who have been diagnosed.

### **Current Report**

This report provides the first attempt to assess *observed* cancer survival of cases held by the Guam Cancer Registry. The initial section (Analysis 1) focuses on lung and bronchus cancer because it is Guam's most common cancer site for both incidence and mortality. This section contains the most informative information. The subsequent section (Analysis 2) reports the results of the mortality analyses for the top five and selected other cancers grouped by sex, ethnicity, cancer site, and age.

## **METHOD**

### **Data Sources & Inclusion Criteria**

All cases were obtained from the Guam Cancer Registry, part of the Surveillance, Epidemiology, and End Results (SEER) program, and a member of the North American Association of Central Cancer Registries (NAACCR). The registry collects cancer cases from all Guam sources as provided by law (passively from all required sources and actively from key sources) and conducts active follow up of Guam cases (see Guam Public Law 24-198; Guam Cancer Registry Policy and Procedures Manual, 2011; and NCI SEER Manual, 1999).

As per my request, the Guam Cancer Registry extracted all cancer case incidences for the top five primary mortality sites between the years 2000 through 2009. The top mortality sites for males were lung, prostate, colorectal, liver, & nasopharyngeal (NPG) and for females were lung, breast, colorectal, cervix, & liver. A few additional sites of potential interest were included (lip, oral cavity, & pharynx; pancreas; hematopoietic; & uterus). The cases were extracted 23 January 2014 by Registry staff. A second data extraction included cancer cases (*uncensored*) deceased between 1 January 2000 and 31 December 2009 (constituting the *mortality* analyses). Later, all survivors as of that date were included and both deceased and alive cases were included in the *survival* analyses. Total Cases included in the final extraction were 1,551 with 56 censored cases (those who survived after the cutoff date).

To facilitate race/ethnicity analyses relevant to the Guamanian population, cases were re-coded to align with the Guam Epidemiological Work Group recommended five classifications (Chamorro, Filipino, Micronesian Non-Chamorro, Asian Non-Filipino, and Caucasian/other). The data set included 1490 cases in these five ethnicities with 635 Chamorro, 325 Filipino, 66 Asian non-Filipino, 61 Micronesian non-Chamorro, and 403 Caucasian/other.

**Data Analysis.** The data set was imported into Microsoft Excel for analyses. The date columns were converted to 'date data' and the number of months of survival was determined by subtracting the diagnosis date from date of death/last follow up. Insufficient date information was provided in 42 cases: 31 cases provided no 'day' of diagnosis and 11 cases included no 'month' of diagnoses. A day of '15' was inserted for the cases lacking day data, while cases with no month of diagnoses information were censored (10 males, 1 female).

## **RESULTS & DISCUSSIONS**

### **Analysis 1: Cancer Survival Results**

**Cases.** The survival analysis limited to cancer originating in the lung and bronchus only. The analysis includes 272 cases with 94 females and 178 males. The ethnic groups considered include 135 Chamorro, 65 Filipino, and 43 Caucasian/other. Micronesian non-Chamorro and Asian non-Filipino were censored due to too few cases (19 and 10 respectively).

**Charts.** The following charts display the survival curves as a percent of the total lung & bronchus cases (y-axis) across five years (x-axis, in months) by sex (Figure 1), ethnicity (Figure 2), and comparison to other datasets and analytic methods (Figure 3).

**Lung & bronchus survival by sex.** Figure 1 displays all Guam lung & bronchus cases combined (middle line), female cases (top line) and male cases (bottom line). Female cases show a higher survival curve than males, which follows national trends (Ries, Young, Keel, Eisner, Lin, & Horner, 2007). The five-year observed survival percentages for all lung & bronchus cases on Guam is 25% with 30% for females and 23% for males (the right end point for each curve).

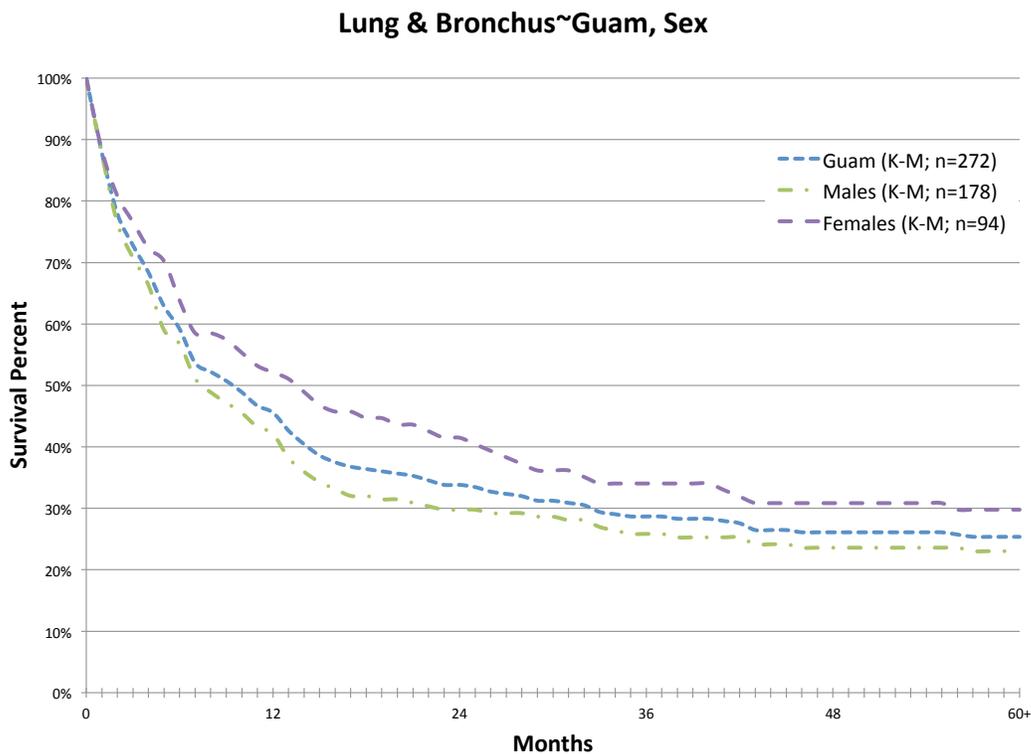
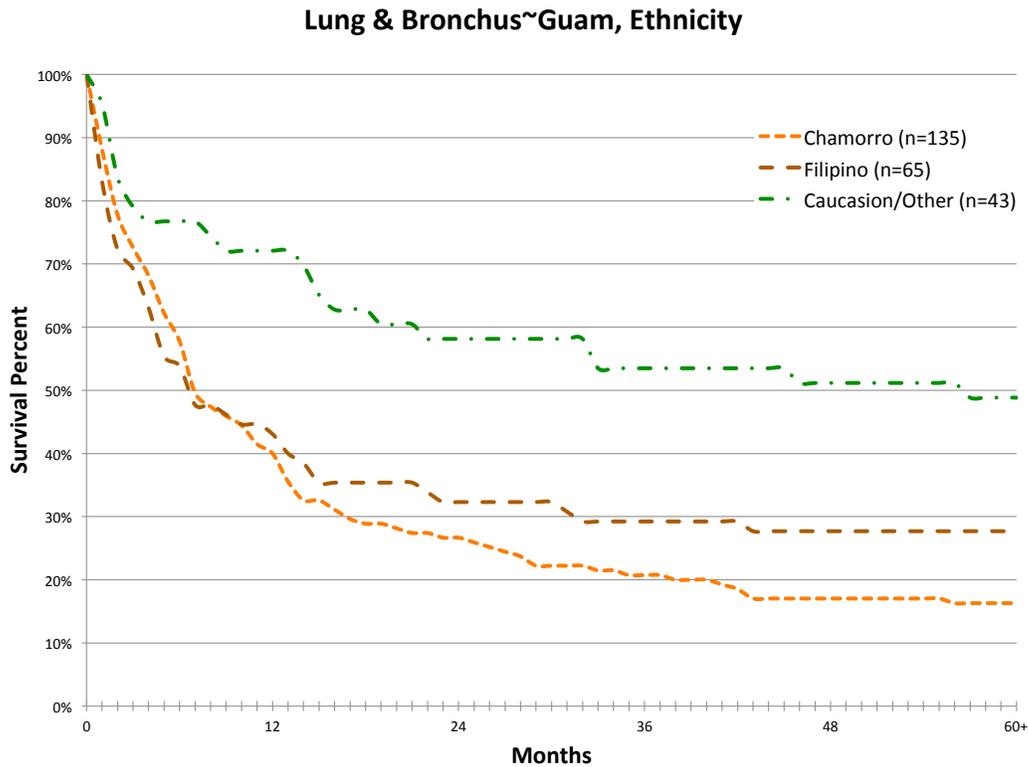


Figure 1: Guam lung & bronchus survival percent at one-year intervals (in months) by sex.

**Lung & bronchus survival by ethnicity.** Figure 2 displays lung and bronchus cases grouped by Chamorro, Filipino, and Caucasian/Other. The striking result is the separation that occurs for the three ethnicities. The Chamorro and Filipino curves track

similarly through the first 12 months (40% and 43% respectively) but then diverge as the Chamorro percent decreases to a five-year survival of 16% while the Filipino five-year survival only decreases to 28%. The Caucasian/Other group remains nearly double the other two ethnicities (one-year is 72%, five-year is 49%). The high survival rate for this group is unusual and likely due to other factors (see Discussion for details).



Ehlert\_CS\_Lung\_2000-2009.xlsx

December 2014

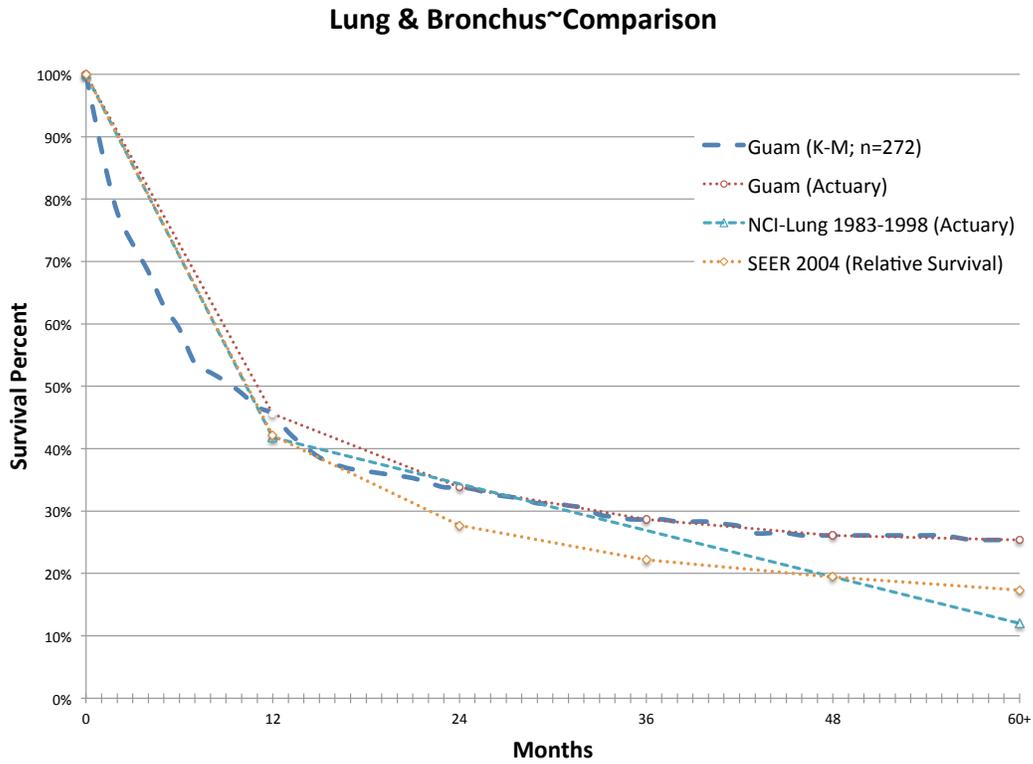
Figure 2: Guam lung & bronchus survival percent at one-year intervals (in months) by three ethnicities.

**Lung & bronchus survival comparisons.** Figure 3 depicts comparisons between (a) analyses of Guam cases using two methods, the Kaplan-Meier and actuary, (b) National Cancer Institute actuarial method, and (c) a *relative* survival analysis of the 2004 SEER national data.

First, consider the two analyses of Guam data. The primary difference between actuary and Kaplan-Meier occurs during the first year. The survival percent estimated using Kaplan-Meier reduces more rapidly than the actuary because the former calculates a point every month. The rapid decline indicates that many cases decrease soon after diagnosis, a result that is masked when using the actuarial method.

Comparing Guam to the national data reveals important differences. The actuarial results for Guam (broken line) and the NCI data (dotted line; from Compton et al., 2012,

Figure 2.2) are similar through 12 months (46% and 42% respectively) but then diverge by five-years (25% and 12% respectively). The national data survival curves using observed (actuary, dashed line) and relative (dotted line) survival rates show differences as well with identical values at one year but differences at five years (observed is 12% and relative is 17%). The limited data points within the five-year interval weaken direct comparisons; however, the different values at five years indicate practical differences.



Ehlert\_CS\_Lung\_2000-2009.xlsx

December 2014

Figure 3: Guam lung & bronchus survival percent at one-year intervals (in months) by calculation method.

### Analysis 1: Survival Analysis Discussion

The Guam Cancer Registry data are essential resources for understanding cancer on Guam. The work of the Registry staff is laudable, at times heroic. The survival analysis, however, revealed some limitations that need to be addressed for a future survival analysis using Guam data to meet national standards.

### Withdrawal Cases

The different survival curves for the three ethnicities are striking. Additional analyses are necessary to confidently explain why the differences exist. At this point, I suspect the differences are artifacts, most likely due to cases lost due to follow up (*withdrawal*) and should be censored. That is, determining if a case is still alive takes additional and sometimes laborious work. Nevertheless, the work is essential to accurately estimate the cancer survival rate. The Guam data may have incomplete end-point information that is different for the ethnic groups.

The concerted effort of the cancer community on Guam over the past decade has substantially improved the collection of cancer incidences, thereby increasing reports of cancer. However, the end point for survival analysis, vital status at a specific time, may be less accurate particularly for individuals who leave island after diagnosis. It is more difficult to determine their status, which would yield results similar to those displayed in Figure 2. The Chamorro cases are most likely complete since those born and raised on Guam are more likely to remain on Guam after diagnosis, return to Guam after treatment, or be buried on Guam. Chamorros constitute the higher proportion of Guam residents and, thus, most likely have the most accurate vital status. The Caucasian/Other cases are least likely to have been born or raised on Guam and, thus, most likely to depart Guam after diagnosis, leave Guam for treatment, and least likely to return to Guam for burial upon death. These cases are less likely to be included in the Guam registry and less likely to have accurate follow-up for Vital Status. The Guam Cancer Registry most likely would not know the true vital status nor receive notification of death. These cases, then, would be classified as 'alive' thereby (artificially) increasing the survival rate similar to that seen in Figure 2. A similar rationale would hold for Filipino cases, since they could easily go to the Philippines for treatment, likely have family there, may not return after treatment, and least likely to return for burial. The ethnic disparity displayed in Figure 2 is most likely due to the inclusion of cases that should be withdrawn.

Additional factors could contribute to the large ethnic disparity as well. Possible factors include stage at diagnosis, the treatment type received and subsequent maintenance, insurance coverage, and socioeconomic issues.

### **Proposed Remedy**

I suggest two solutions on how to handle the withdrawal cases. In the long run, the community is served best if all residents are treated equally. Therefore, more effort should be devoted to assuring accurate and complete information for all cases regardless of possible challenges related to ethnic classification. Another solution presents itself to address immediate concerns. I suggest a rule be developed to censor cases that leave island or have no follow up information after a specified time (the SEER program may have suggestions). Perhaps the Guam Cancer Control Coalition could assist with developing the rule. The first action, however, is to determine if withdrawals explains the disparity, which requires additional research.

A procedural change also could help identify withdrawal cases. The registry could include a classification that differentiates between 'last contact' and 'confirmed deaths.' The current practice is to record them in a single column, making it impossible to disambiguate last contact date from death date.

### **Adjustments**

This analysis focused on observed survival rates without adjustments. However, age and cause of death adjustments likely would improve survival estimates. Age adjustment may be important particularly if Guamanians get cancer younger than others. Any effect on survival of an earlier onset of cancer incidence should be more evident when the cases are normalized by age. Determining relative survival by adjusting the observed survival by cause of death also should improve the survival estimates. However, Guam's small numbers may challenge the interpretation. That this analysis used observed survival *only* could explain some or most of the differences between the Guam data and the national data presented in Figure 3. Comparisons to other populations must be made cautiously without these standard adjustments.

### **Analysis 2: Mortality Analysis Results**

The mortality data are reported in four clusters. These analyses include only deceased cases recorded before January 2010. The first analysis (2a) considered all cases in the first data extraction (see Figures 4, 5, & 6) grouped by sex. The second analysis (2b) focused on the first year after diagnosis (see Figures 7, 8, & 9). The third analysis (2c) grouped results by cancer origin site (see Figures 10 & 11) and ethnicity (see Figures 12 & 13). The final analysis (2d) reanalyzed the dataset excluding all cases that survived less than one month after diagnosis (see figures 14 through 22). Figures 21 and 22 report all included cancers by age at diagnosis. Each analysis includes a chart depicting raw count (first) and proportion of total (second).

#### **Analysis 2a: Mortality curves by sex**

The 64% reduction in the first year and the 5% five-year survival depicted in Figure 4 are troubling. This result is primarily due to limiting the data set to deceased cases prior to the end of 2009. In effect, this analysis excluded all persons who remained alive at the end of the time frame thereby preventing a *survival* analysis. The next two charts show that although males have a higher incidence (see Figure 2) the mortality rate is about the same (see Figure 3).

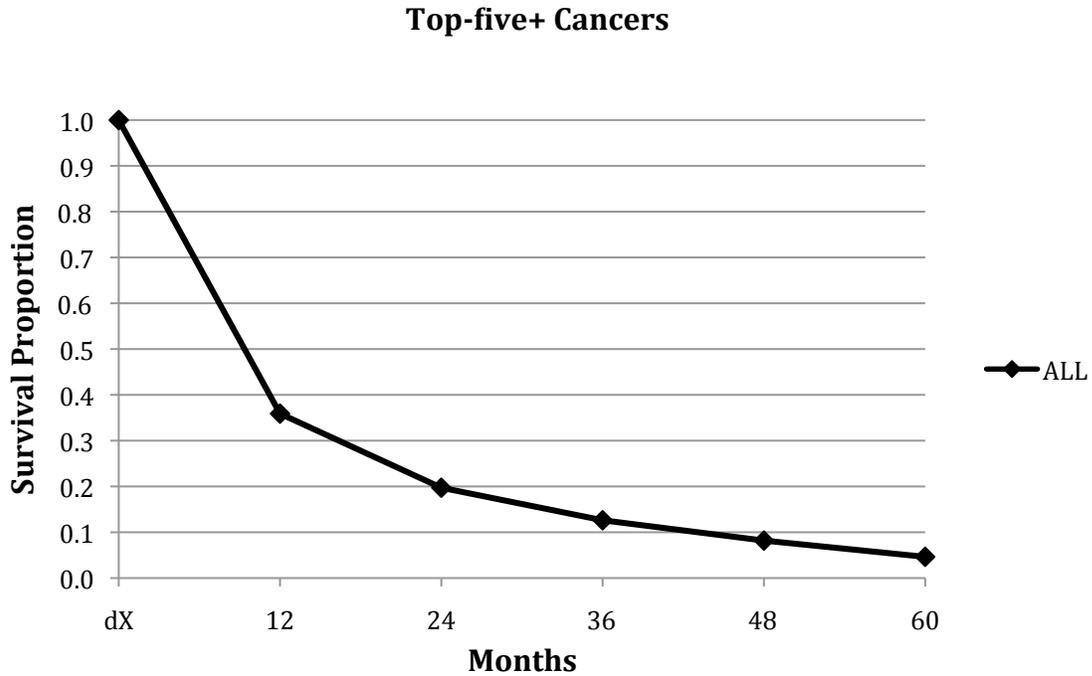


Figure 4. All ethnicities; survival proportion at one-year intervals (in months)

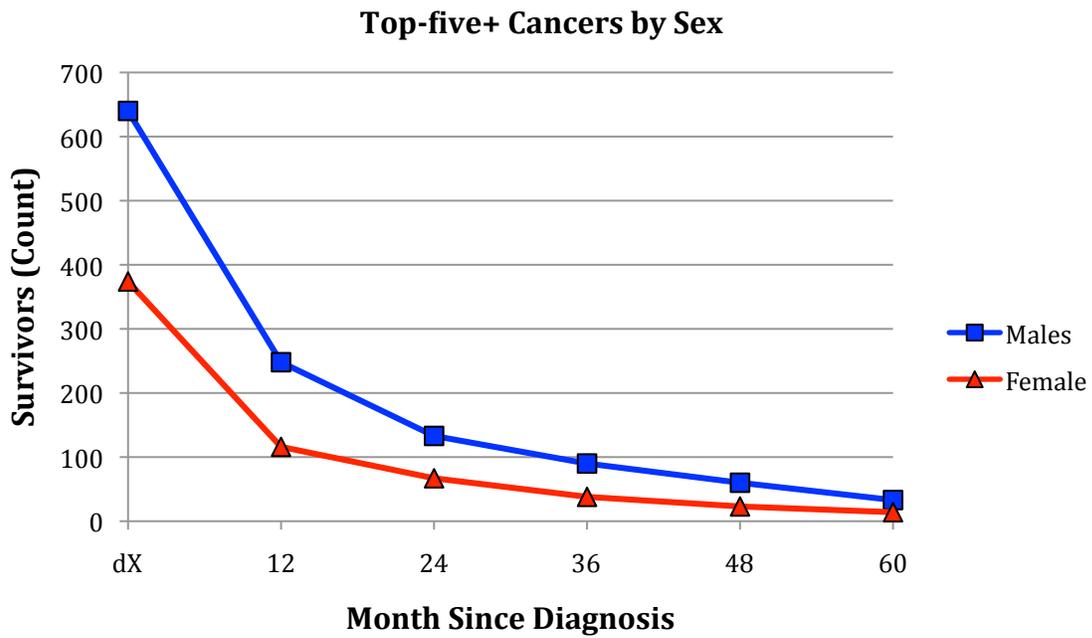


Figure 5. All ethnicities; survival count at one-year intervals (in months)

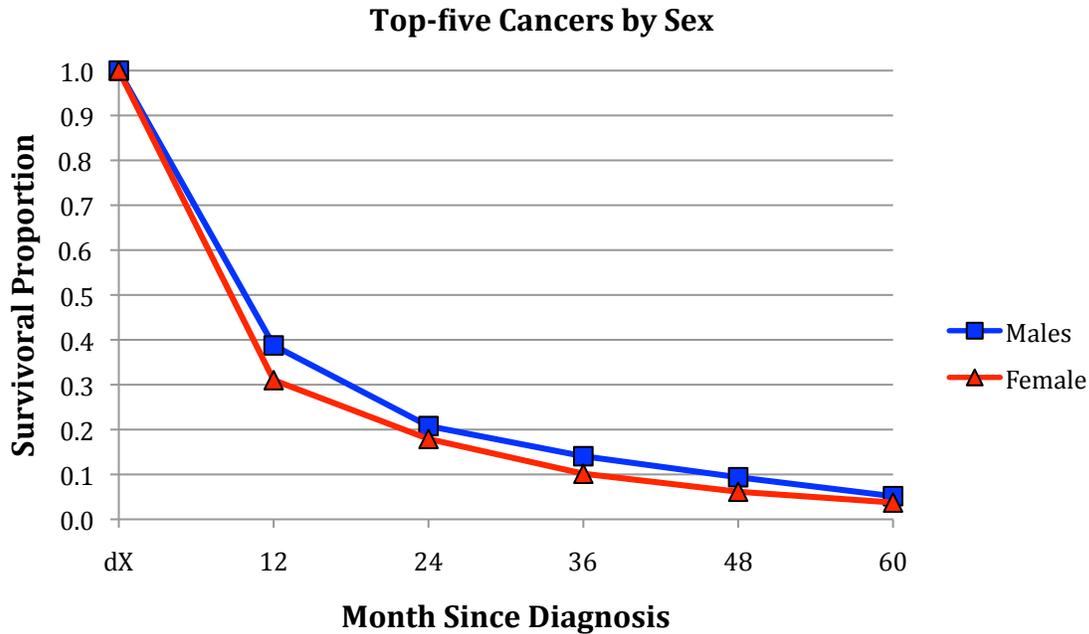


Figure 6. All ethnicities; survival proportion at one-year intervals (in months)

### Analysis 2b: Mortality curves for first-year survival by sex

The death certificate initiated (DCI) cases (n= 174) and one-month survival cases (n= 103) are substantial. Registry staff obtained additional information on some of these cases but many are death certificate only (DCO). Standard practice is to include DCO cases (see Parkin & Hakulinen, 1991, and Hanai & Fujimoto, 1985) based on the assumption that they are an insignificant proportion of the data set. We may want to reconsider this practice for Guam, or better understand why the DCI and DCO proportions are so high.

In order to investigate the substantial loss during the first year, I plotted the first-year survival function by month, with *Month Zero* defined as the date of diagnosis (dX) and *Month 11* as completing *through* the 11<sup>th</sup> month (one-year post diagnosis). These data indicate that nearly 30% of cancer patients expire within the first month after diagnosis. These patients likely are diagnosed near the end of life with effectively no chance for treatment.

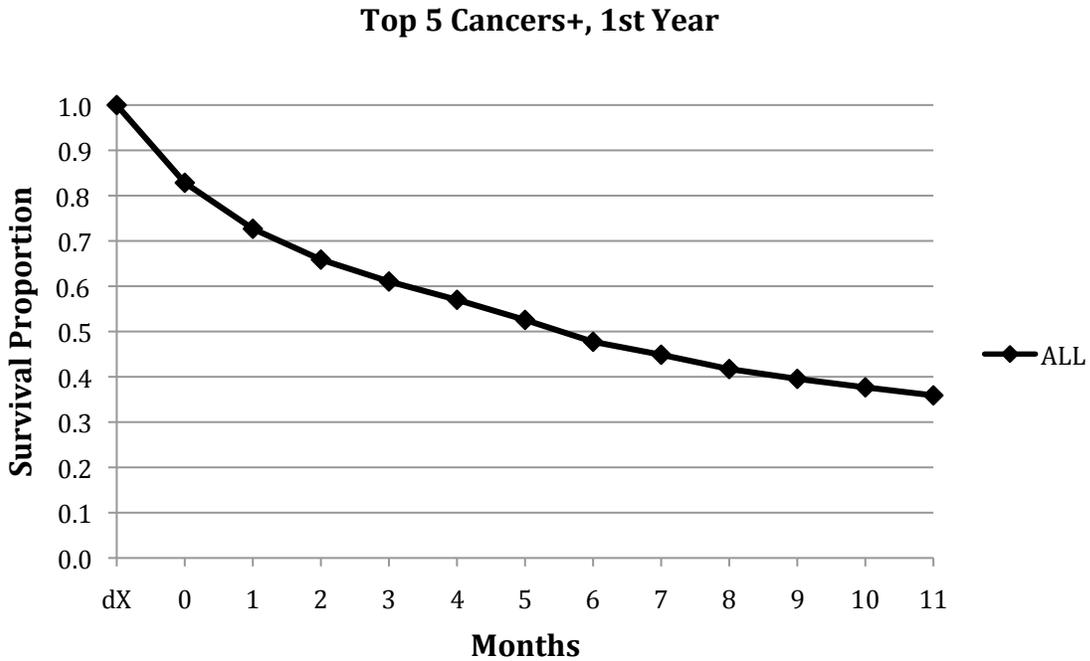


Figure 7. All ethnicities; first year survival proportion at one-month intervals

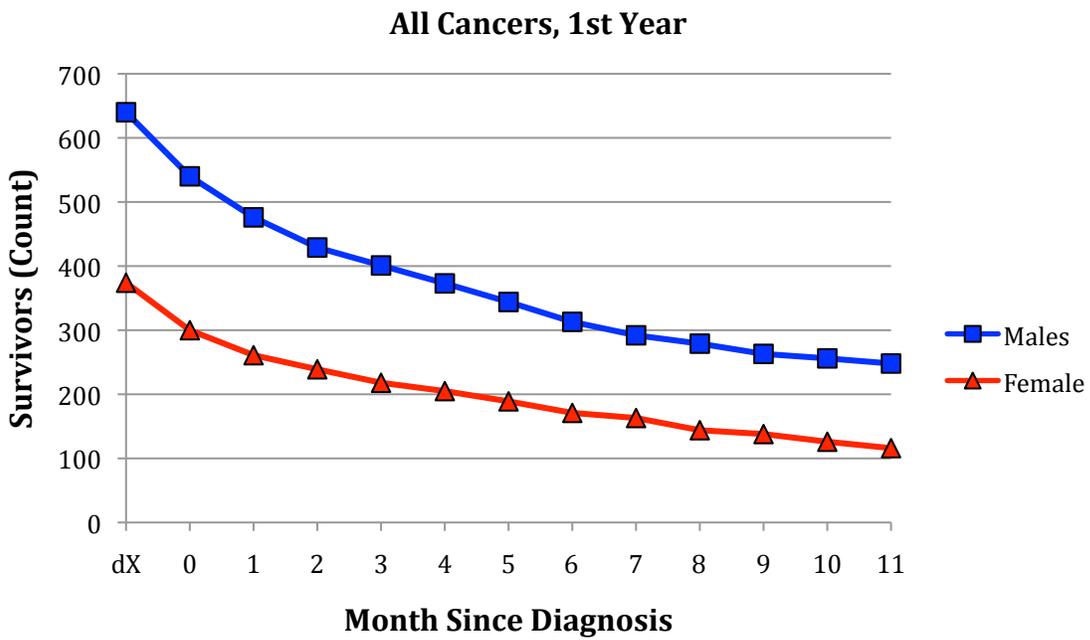


Figure 8. All ethnicities by sex; first year survival count at one-month intervals

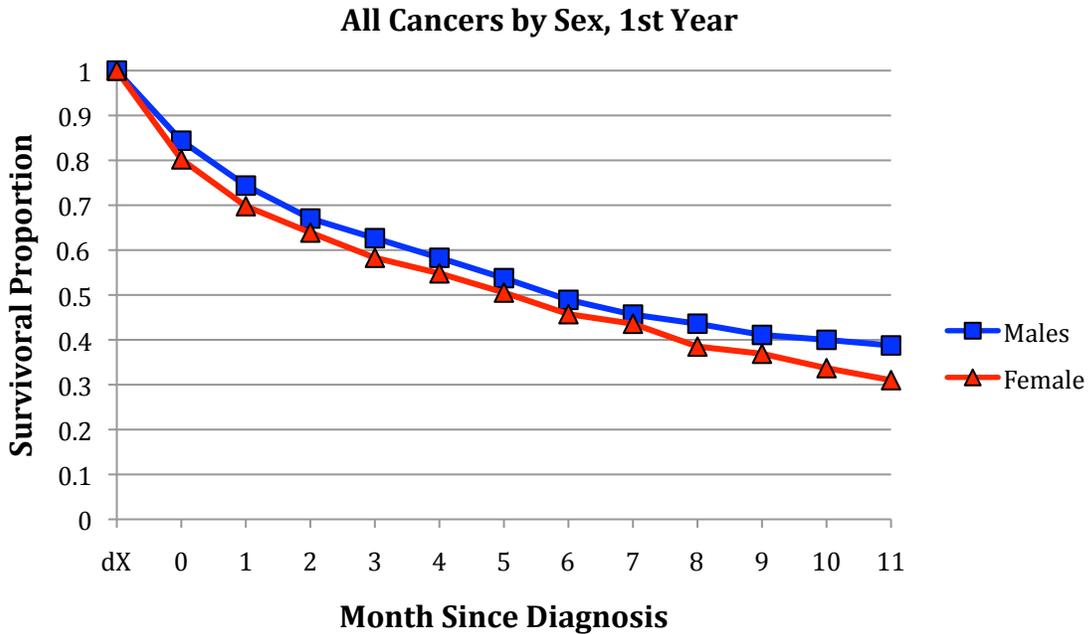


Figure 9. All ethnicities by sex; first year survival proportion at one-month intervals

**Analysis 2c: Five-year survival curves by (a) site and (b) ethnicity.**

Additional analyses were conducted by primary cancer site and ethnicity. The site analysis (n=791) includes both sexes for colorectal, liver, and lung & bronchial. Prostate and breast are included as among the top five for each sex. The ethnicity analysis (n=871) includes Chamorro, Filipino, and Caucasian (White) only.

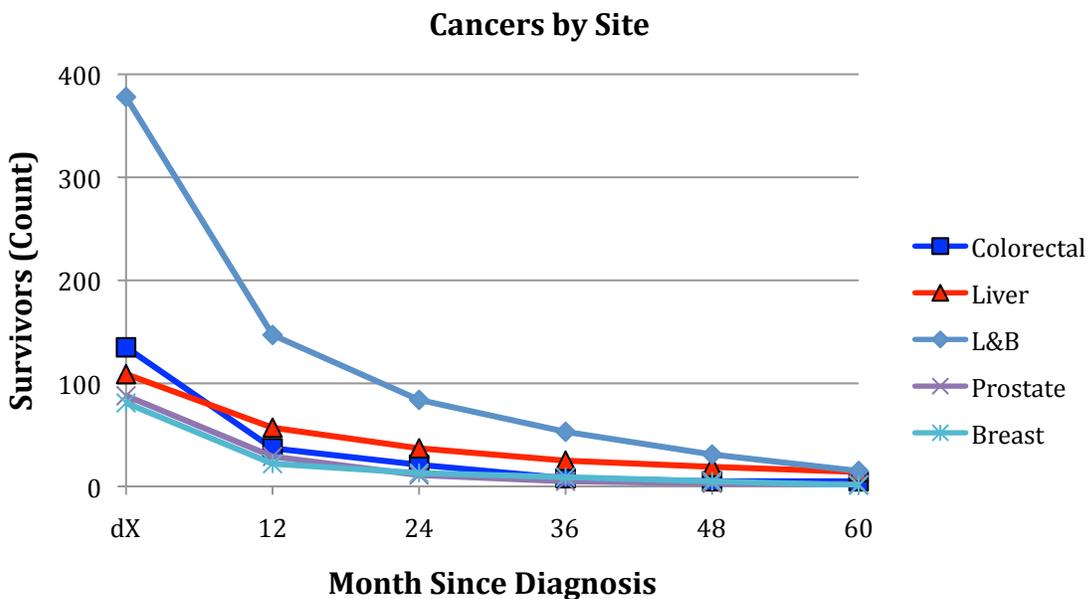


Figure 10. Top 5 cancers by site; survival count at one-year intervals (in months)

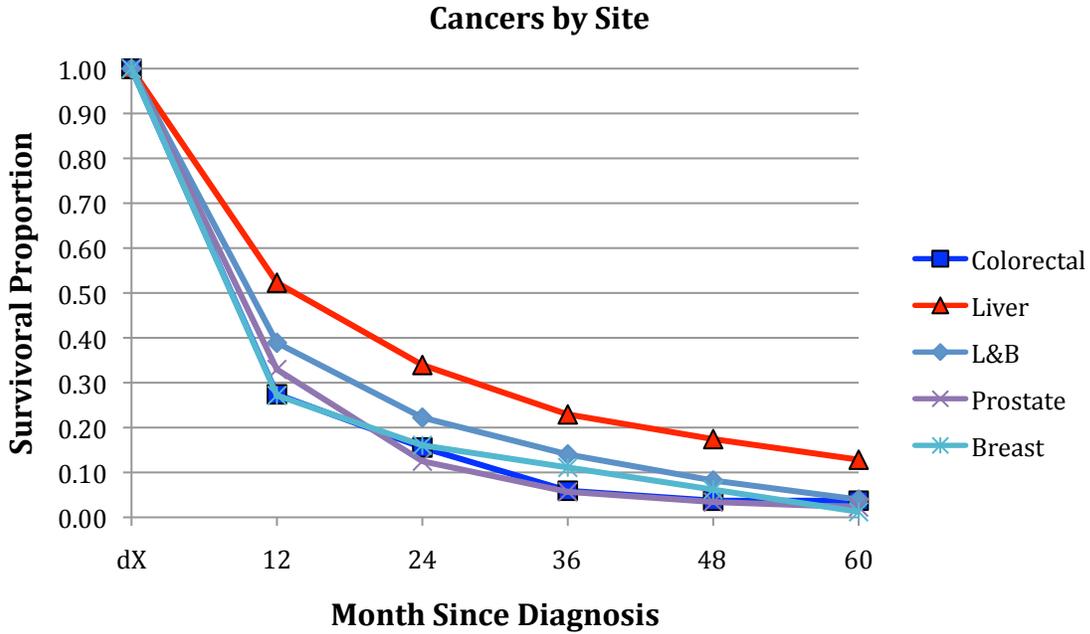


Figure 11. Top 5 cancers by site; survival proportion at one-year intervals (in months)

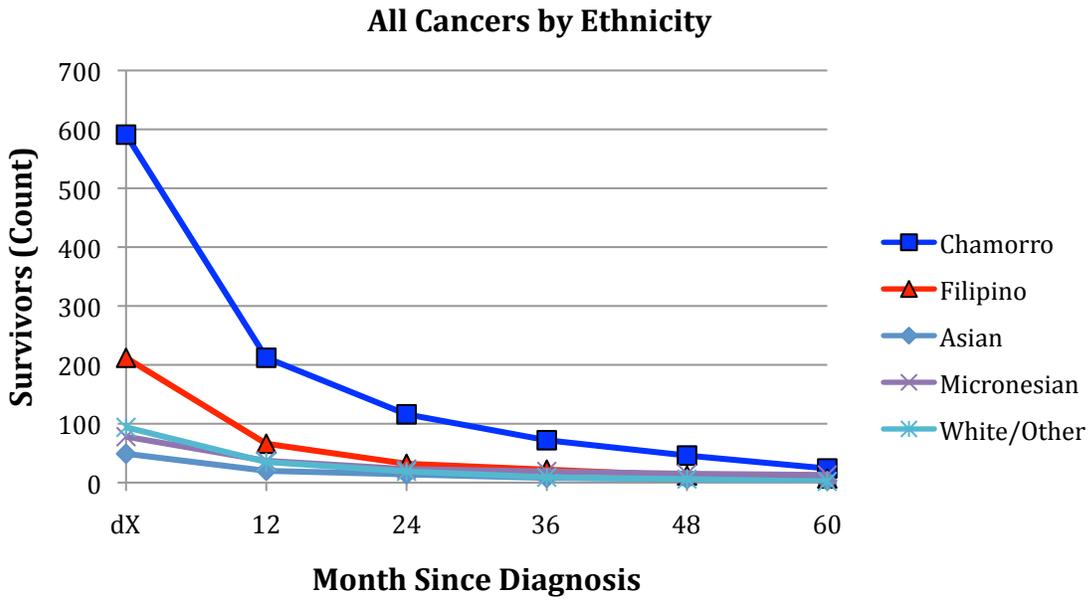


Figure 12. Cancers by ethnicities; survival count at one-year intervals (in months)

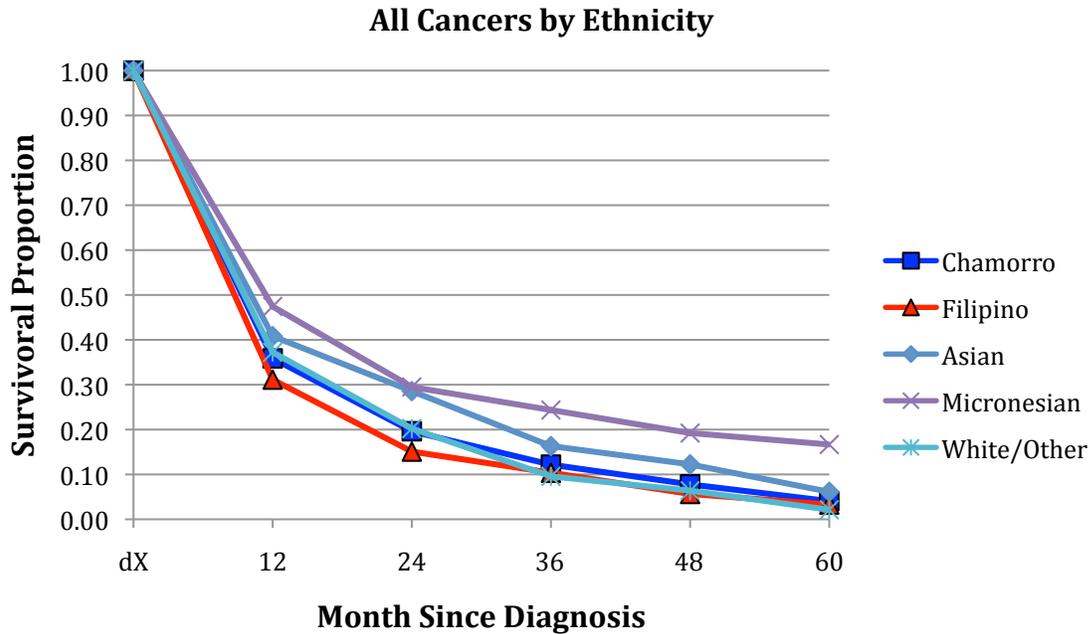


Figure 13. Cancers by ethnicity; survival proportion at one-year intervals (in months)

#### **Analysis 2d: Excluding the 'zero' durations.**

There is a question of the effect of including cases with the survival duration 'zero.' The following section reports the main analyses redone excluding the 'zero' survival cases and the addition of an age analysis. In these analyses, a 'zero' occurred for a case that survived any duration less than the first month. That is, month equals zero when death is recorded as less than the next month after diagnosis. We may wish to distinguish between when the diagnosis date equals death date from those cases that survive a few days/weeks but less than a month. The present analyses include the following sample sizes: All and Ethnicity, n=849, and Site, n=666.

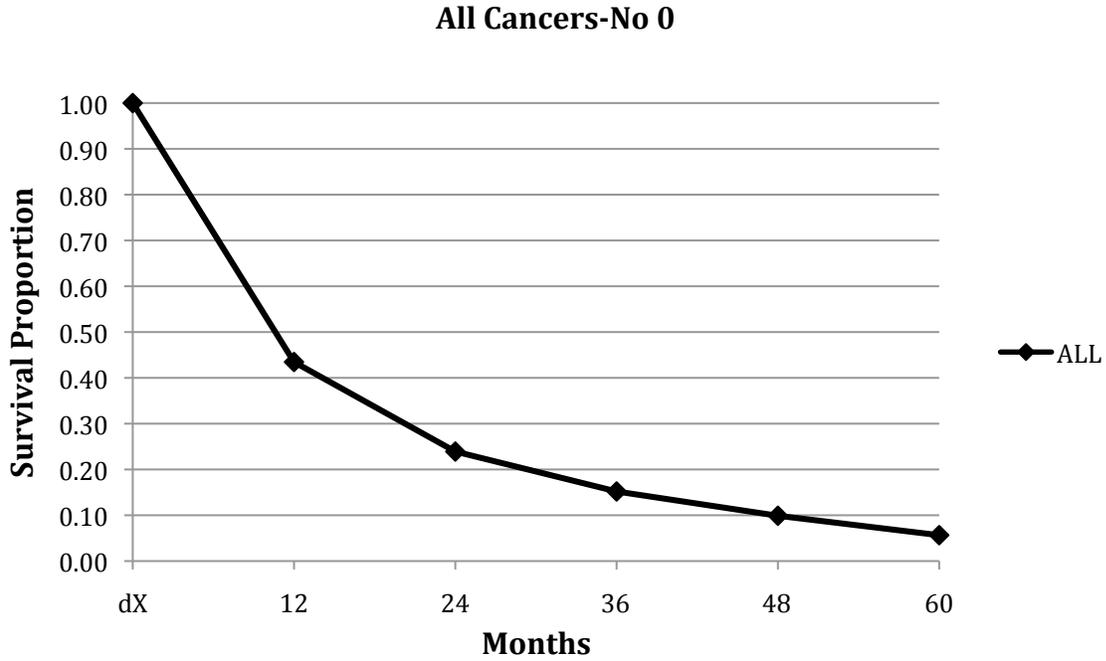


Figure 14. All ethnicities; survival proportion at one-year intervals (in months); 'zero' cases excluded

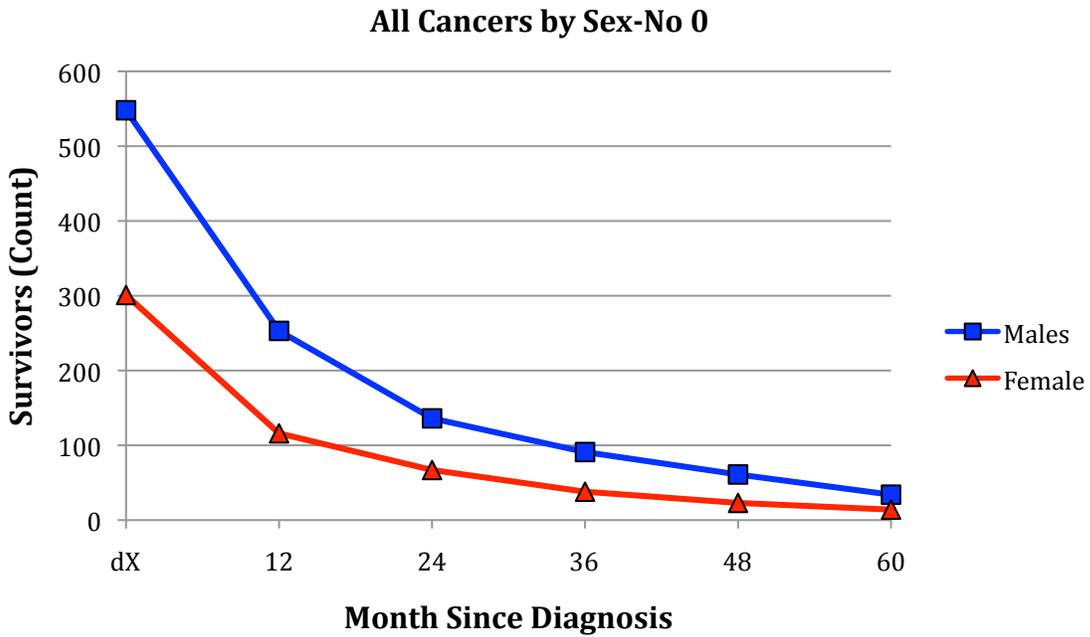


Figure 15. All ethnicities by sex; survival count at one-year intervals (in months); 'zero' cases excluded

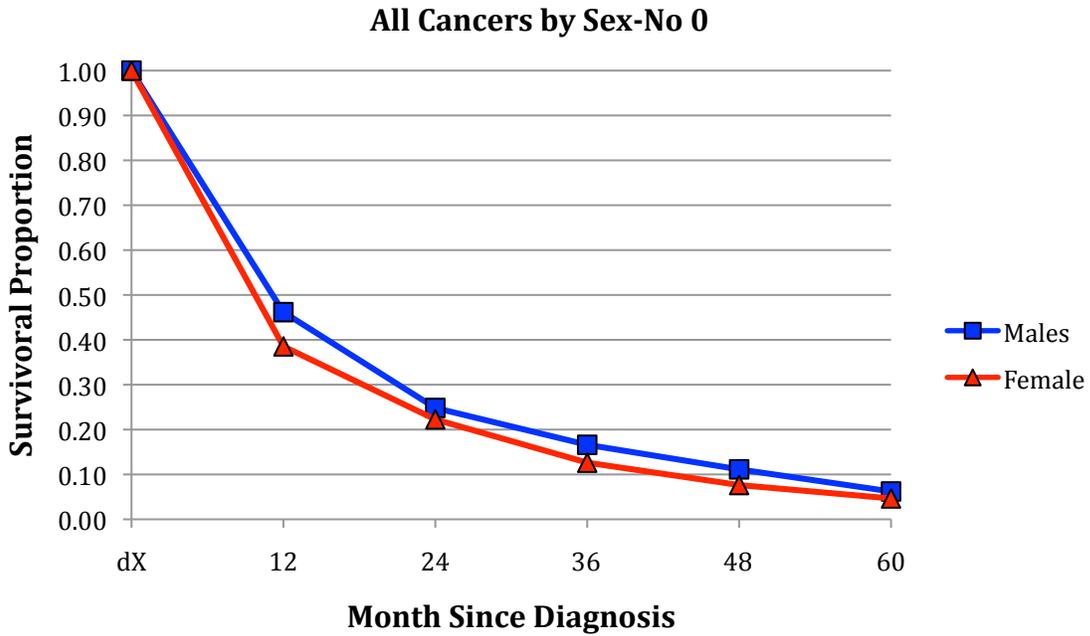


Figure 16. All ethnicities by sex; survival proportion at one-year intervals (in months); 'zero' cases excluded

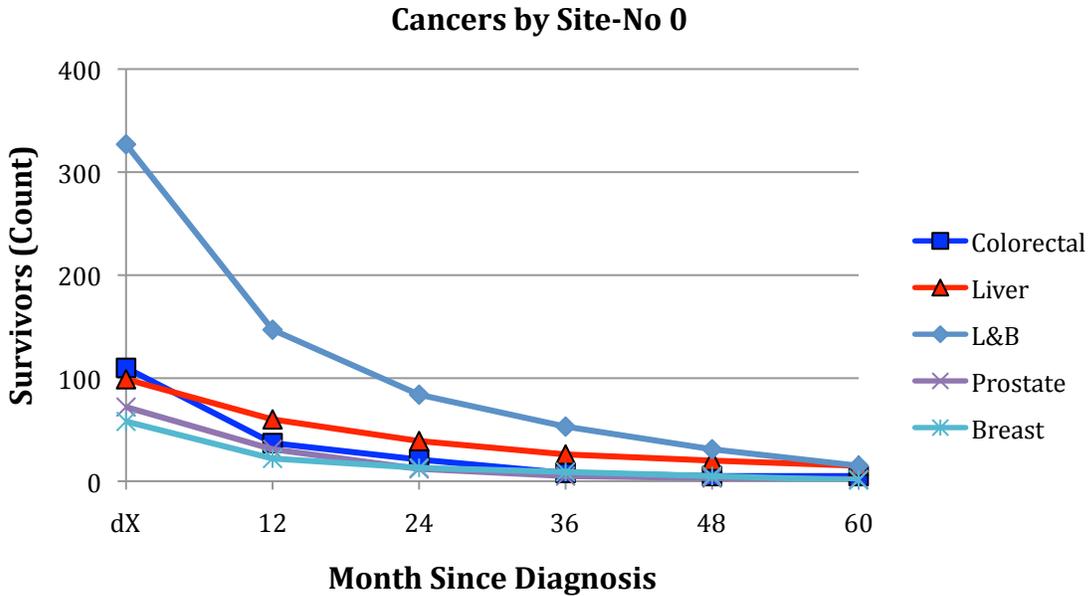


Figure 17. Top 5 cancers by site; survival count at one-year intervals (in months); 'zero' cases excluded

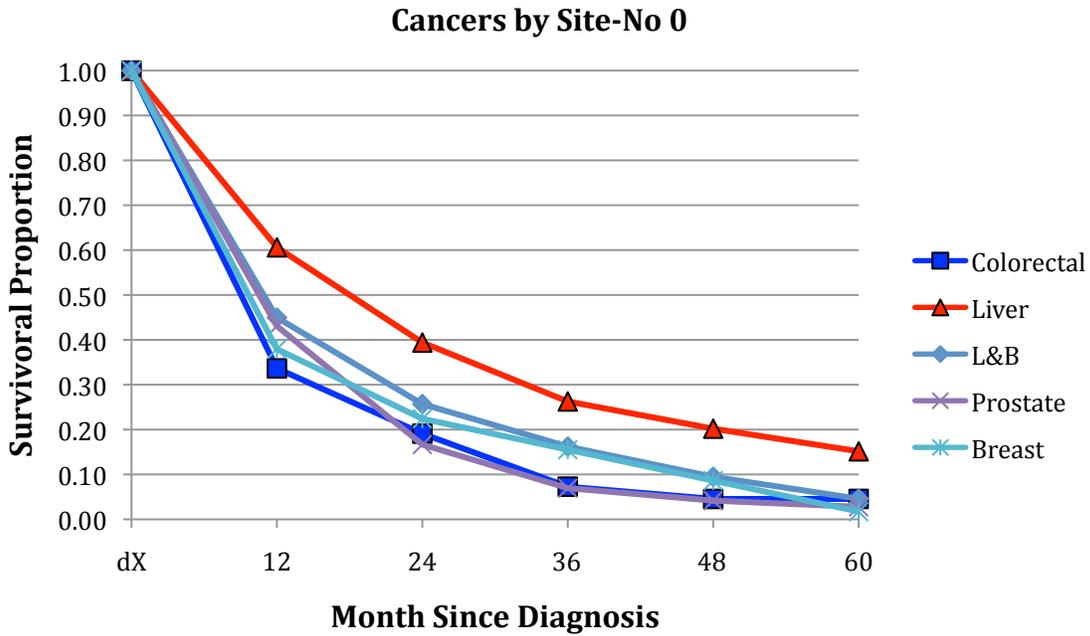


Figure 18. Top 5 cancers by site; survival proportion at one-year intervals (in months); 'zero' cases excluded

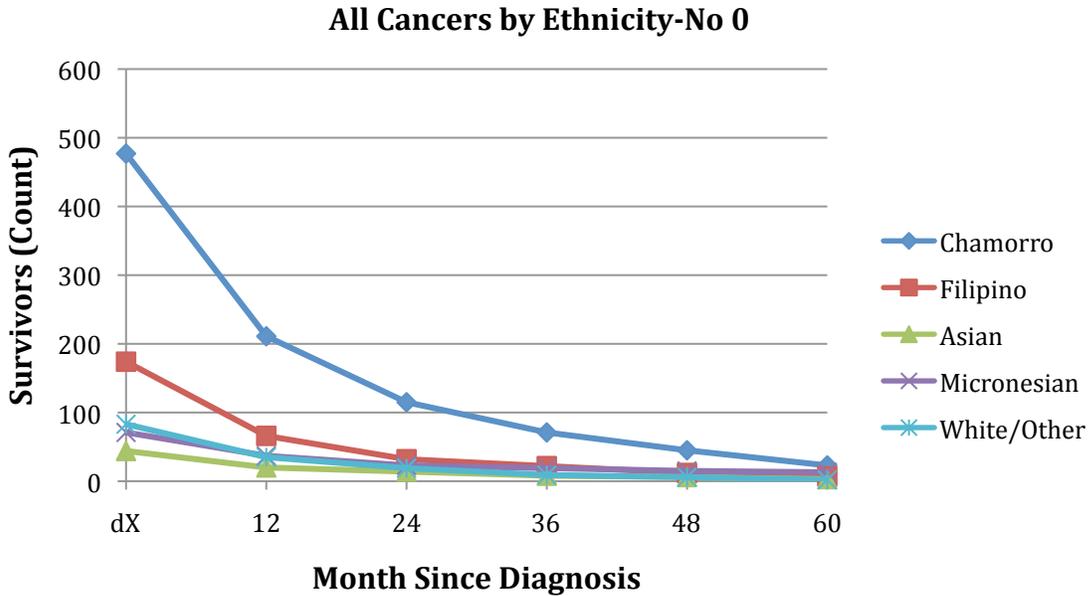


Figure 19. Cancers by ethnicity; survival count at one-year intervals (in months); 'zero' cases excluded

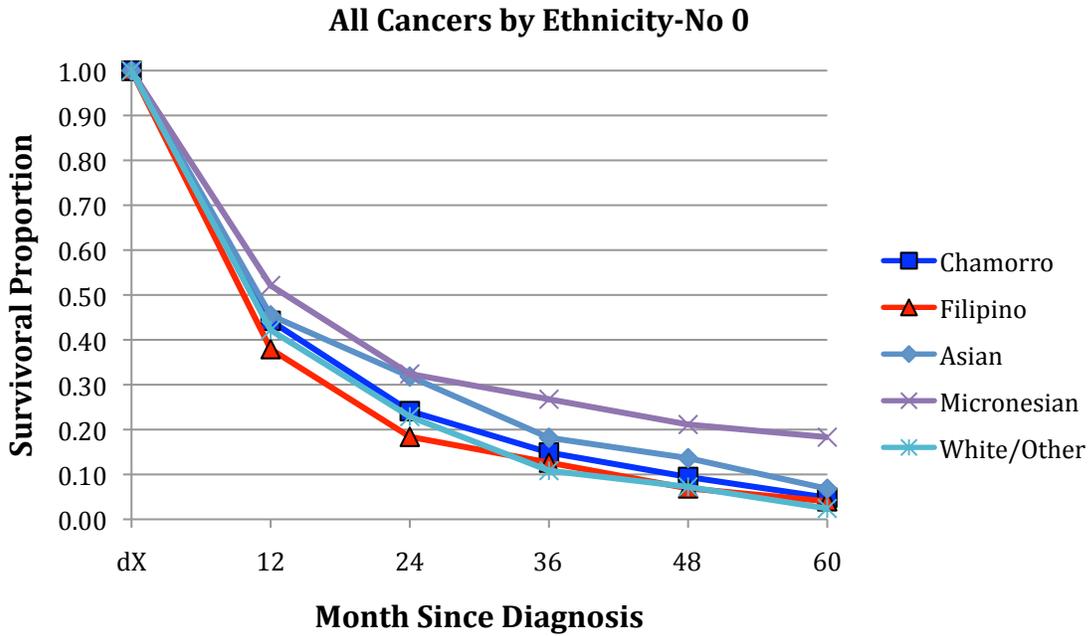


Figure 20. Cancers by ethnicity; survival proportion at one-year intervals (in months); 'zero' cases excluded

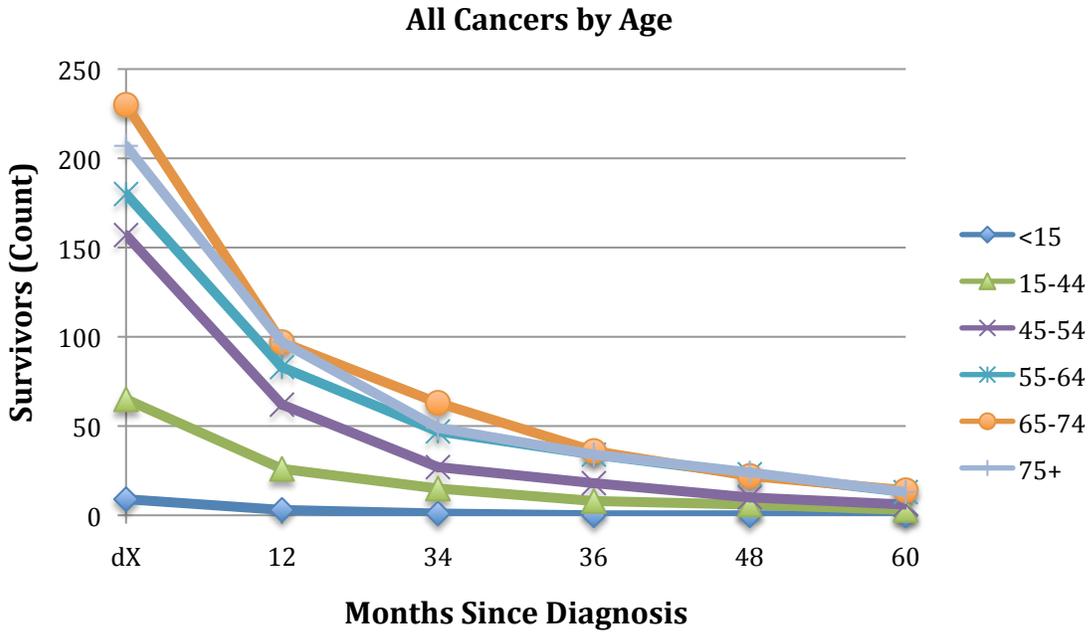


Figure 21. Cancers by age; survival count at one-year intervals (in months); 'zero' cases excluded

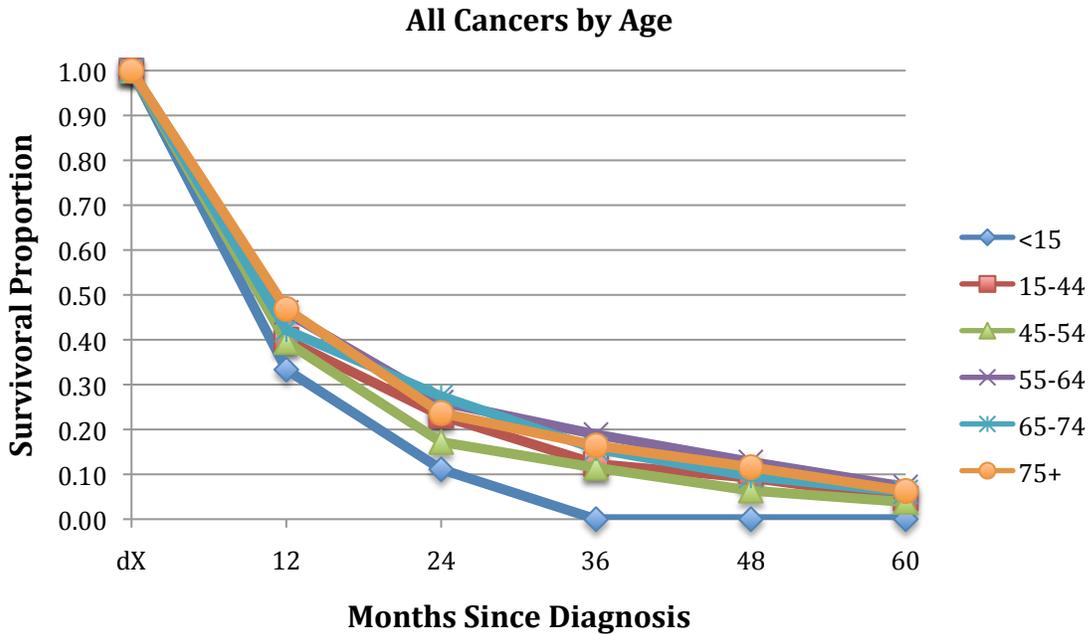


Figure 22. Cancers by age; survival proportion at one-year intervals (in months); 'zero' cases excluded

### Analyses 2: Mortality Analysis Discussion

The mortality analyses reveal that the vast majority of deceased cases died within the first year post diagnosis. Careful evaluation by interested persons could guide future research. It is difficult, however, to draw firm conclusions from these results because the analysis excluded cases alive after 2009.

### CONCLUSION

Conducting this research was a challenge; yet much was learned that could improve future research. Of particular interest is the probable effect of withdrawal cases. The effect of cases lost to follow-up ('withdrawal') must be determined to even hope for valid comparisons with national and international data sets. The Guam Cancer Registry could improve its assessment of potential withdrawal cases possibly by creating a rule to censor any cases without contact during some specified period. The proper time period would depend on how regularly the complete Guam data set could reasonably be evaluated by the staff, which is acknowledged as an extremely difficult task given the limited staff and its current responsibilities. A longer no-contact duration than may be the standard would be a more conservative measure for censoring.

Future research could focus, initially, on determining the effect of withdrawals on the disparity between the ethnicities. Following that, survival analyses should proceed to other cancer sites, specific ages, and socioeconomic indicators. Eventually, one would hope to compare Guam's data to equitable data from other locales. Usually this means to the US national data. However, future analyses might target regions with high concentrations of Chamorro and other Pacific Island ethnicities rather than national averages.

## REFERENCES

- Compton, C.C. et al. (2012). *AJCC Cancer Staging Atlas: A Companion to the Seventh Editions of the AJCC Cancer Staging Manual and Handbook*. DOI 10.1007/978-1-4614-2080-4\_2
- Cho, H., Howlader, N., Mariotto, A.B., & Cronin, K.A. (2011). Estimating relative survival for cancer patients from the SEER Program using expected rates based on Ederer I versus Ederer II method. *Surveillance Research Program, National Cancer Institute, Technical Report #2011-01*
- Hanai, A. & Fujimoto, I. (1985). Survival rate as an index in evaluating cancer control. In, Parkin, Wagner, & Muir (eds.), *The Role of the Registry in Cancer Control* (IARC Scientific Publication No. 66). Lyon, International Agency for Research on Cancer.
- NCI SEER Program (1999). *Self Instructional Manual For Cancer Registrars, Book 1: Objectives and Functions of Cancer Registries, Third Edition*.
- Parkin, D.M. & Hakulinen, T. (1991). Analysis of survival. Chapter 12 in Jensen, Parkin, MacLennan, Muir, & Skeet (eds.) *Cancer Registration, Principles and Methods*. IARC Scientific Publications No. 95. IARC Press, Lyon, pp 159–176.
- Ries, L.A.G., Young, J.L., Keel, G.E., Eisner, M.P., Lin, Y.D., Horner, M-J. (2007). SEER Survival Monograph: Cancer Survival Among Adults: U.S. SEER Program, 1988-2001, Patient and Tumor Characteristics. *National Cancer Institute, SEER Program, NIH Pub. No. 07-6215, Bethesda, MD*.